A selected bibliography on parameter optimization methods suitable for hybrid computation

by ELMER G. GILBERT
The University of Michigan

FOREWORD

A common situation in engineering and science is the following: given a dynamical system whose motion depends in some way on a set of parameters, find admissible values for the parameters such that a numerical index of system performance (or cost) is maximized (minimized). Since such *parameter optimization* problems are generally impossible to solve analytically, computer methods involving successive systematic evaluations of the performance index are appropriate. Because many evaluations (each requiring a complete solution of a system of differential equations) are usually necessary, the hybrid computer, which combines high-speed solution of differential equations with flexible programmed control and rapid numerical calculations, is well adapted to this type of computation. This is substantiated by the appearance of a number of papers which treat optimization methods and devices specifically oriented toward hybrid computation.

The bibliography which follows lists the more important of these papers and in addition includes references which contain applicable mathematical theory, promising computational methods which have not yet been used in hybrid computation, representative examples of work in related fields, and sources of additional references. Since work on optimization has appeared in such diverse fields as operations research, statistics, process control, adaptive control, and the response of human operators (references 1, 2, A1, B1, B3, B6, C2, E1, S1, W2), no claim to comprehensiveness is made. Those items which appear to be most readable, representative, and cogent are marked by an asterisk (*). The article by Bekey[B6] is an introduction and general survey of hybrid computer optimization. General treatments of optimization methods appear in references B5, H2, K2, L1, S5, and W2. We now give a brief guide to the bibliography.

The basic mathematical problem which must be solved in parameter optimization is the maximization (or minimization) of a function $F(x_1, x_2, \ldots, x_n)$ of many variables $x_1, x_2, \ldots, x_n$ (the parameters). Standard references, e.g., references K1 (sections 12-15 and 12-16), H2, and L1 (chapter 1), discuss pertinent mathematical questions such as the precise definition of various types of maxima, conditions for the existence of a maximum, necessary or sufficient conditions for a maximum, etc. When equality constraints, $G_i(x_1, x_2, \ldots, x_n) = 0$, $i = 1, 2, \ldots, p$, and inequality constraints, $H_i(x_1, x_2, \ldots, x_n) \leq 0$, $i = 1, 2, \ldots, q$, are imposed on the parameters, the situation is considerably more complex and the corresponding results are to be found in the literature of mathematical programming (references H2, S5, and Z1).

A great variety of methods for numerically solving the maximization problem are available. For our purposes they may be divided into two broad categories: *search methods,* which require the repeated evaluation of F, and *gradient methods,* which in addition require the repeated direct evaluation of the n partial derivatives of F. Search methods have the distinct advantage that the analog portion of the hybrid computer need mechanize only the equations of motion and the calculation of the index of performance. When gradient methods are used, extra analog equipment is needed for evaluating the partial derivatives of F. In particular, "sensitivity functions" or "parameter influence coefficients" must be generated through the solution of certain auxiliary differential equations (references B2, B4, M1, M3, and T1).

Let us consider search methods first. The simplest approach is the exhaustive evaluation of F over a fixed grid of points in the parameter space[W2], a method which requires an unrealistically large number of evaluations if n is greater than two or three. Alternatively, the grid may be continually modified as evaluations of F are made. Such sequential search procedures[W2] are quite efficient for $n = 1$, but become involved and rather inefficient when extended to larger n.

A whole family of methods is based on successively perturbing the vector $x = (x_1, x_2, \ldots, x_n)$ from a base point and then making a change in the base point on the basis of the resulting values of F. The simplest of these is univariate search[B3,W2] where the x_i are perturbed one after the other in a cyclic fashion and where the base point is changed if an improvement in the value of F is obtained. Fewer evaluations generally result if, instead, the partial derivatives (gradient vector) of F are estimated by perturbing $x_1, x_2, \ldots, x_n$ from the base point and then moving in the direction of the gradient from the base point (references F1, S4, W1, W2). The distance moved in the direction of the gradient may be fixed, or may be the result of a sequence of steps, where the goal of the sequence is to maximize F in the direction of the gradient. All these methods are made more complex if constraints are present. They may also get "stuck" on a ridge[N1,W2] or given very slow convergence.

More sophisticated methods of search are generally preferred. Hooke and Jeeves[H1] extended the above approaches by combining a pattern move with the direction-determining (or explore) moves. Practical applications of the Hooke and Jeeves method have shown it to be reliable and quite efficient. References G1 and K3 extend, with considerable complication, the method to constrained optimization problems. Other search algorithms which are more involved than reference H1 are described in references B7, F5, N2, P2, P3, and R2. Rosenbrock's method[R2] has been popular and is somewhat more efficient (in terms of the number of evaluations of F) than reference H1. The methods of references B7 and N2 are little more complex to program and appear promising. The methods given in references P2 and P3 are related to conjugate gradient procedures and, while relatively involved, have very rapid (quadratic) convergence near the optimum.

It was suggested in reference A1 and subsequently that search methods in which the perturbations from a base point are chosen randomly might have advantages[F2,M2,R3,R4,S4,W2]. For example, the computer mechanization of reference M5 is very simple. Although present evidence indicates that random search methods are quite inefficient (when rated on the number of function evaluations for a solution), reference R2 shows that random search may (on the average) be more efficient than certain forms of deterministic search when n is large. Some preliminary work by the writer shows that certain versions of random search are nearly as efficient as reference H1, and have advantages in hybrid computer mechanization.

A variety of gradient methods have been known and used for many years[C1,S4,S5]. They make use of successive corrections of x in the direction of the gradient, and differ mainly in how the distance or step size of the correction is determined. Convergence speed is dependent on the scale factors and coupling of the parameters and, if F is poorly conditioned, *i.e.,* has a "narrow ridge," there is considerable zig-zagging and convergence is slow. A number of schemes which greatly speed convergence have been proposed[F3,F4,P1,S6]. Although these methods tend to be sensitive to computer errors, they are more suitable for hybrid computation than other rapidly convergent methods, *e.g.,* Newton-Raphson[S5], which require the $^1/_2n(n+1)$ second partial derivatives of F. When constraints are imposed, considerably more involved methods in mathematical programming must be applied[H2,R1,Z1].

A number of computer programs and devices have been devised to implement some of the above methods. The Russian literature (references 1, 2 (surveys), F1, N1, S1, S2, S3) describes a variety of search devices or parameter optimizers utilizing analog storage and relay logic which have been used in connection with analog computers and process control. Examples of search mechanizations on hybrid computers are references M5 and W1. Bekey[B6] and Bekey and McGhee[B5] describe computer programs for implementing gradient methods. Little is really known about the efficacy of various search methods in hybrid computation. For some numerical results on several algebraic functions see references B7, F5, N2, P2, and P3.

BIBLIOGRAPHY

1 *Automation and remote control* (ISA translation)
The development of automatic control theory at the institute of automation and remote control 1939-1964
Vol 25 June 1964 pp 691-726

2 *Automation and remote control* (ISA translation)
The development of design principles of automatic control systems and devices at the institute of automation and remote control 1939-1964
Vol 25 June 1964 pp 727-764

A1 R L ANDERSON
Recent advances in finding best operating conditions
Journal American Statistical Association
Vol 48 December 1953 pp 789-798

B1 G E P BOX K B WILSON
On experimental attainment of optimum conditions
Journal Royal Statistical Society Series B vol 13 1951 pp 1-45

B2 L BUSH P ORLANDO
A perturbation technique for analog computers
IRE Transactions EC-8 1959 pp 218-221

B3 S H BROOKS
A comparison of maximum seeking methods
Operations Research vol 7 July 1959 pp 430-457

B4 W BRUNNER
An iteration procedure for parametric model building and boundary value problems
Proceedings of Western Joint Computer Conference
1961 pp 519-533

*B5 A V BALAKRISHNAN L W NEUSTADT editors
Computing methods in optimization problems
Academic Press 1964

*B6 G A BEKEY
Optimization of multiparameter systems by hybrid computer techniques
SIMULATION vol 3 February 1964 pp 19-32 part I and March 1964 pp 21-29 part II

B7 M J BOX
A new method of constrained optimization and a comparison with other methods
Computer Journal vol 8 April 1965 pp 42-52

C1 J B CROCKETT H CHERNOFF
Gradient methods of maximization
Pacific Journal of Mathematics vol 5 1955 pp 33-50

C2 W G COCHRAN G M COX
Experimental designs
Second Edition Wiley 1957

E1 P EYKHOFF
Process-parameter estimation
Progress in Control Engineering
vol 2 Heywood and Company London 1963

F1 A A FEL'DBAUM
Automatic optimalizer
Automation and Remote Control (ISA translation)
vol 19 August 1958 pp 718-728

F2 R R FAVREAU R FRANKS
Random optimization by analog techniques
Proceedings Second International Analogue Computation Meeting Strasbourg
Presses Academiques Europeenes Brussels 1959 pp 437-443

F3 R FLETCHER M J D POWELL
A rapidly convergent descent method for minimization
Computer Journal vol 6 July 1963 pp 163-168

F4 R FLETCHER C M REEVES
Function minimization by conjugate gradients
Computer Journal vol 7 July 1964 pp 149-154

*F5 R FLETCHER
Function minimization without evaluating derivatives
Computer Journal vol 8 April 1965 pp 33-41

G1 H GLASS L COOPER
Sequential search: A method for solving constrained optimization problems
Journal of the Association for Computing Machinery
vol 12 January 1965 pp 71-82

*H1 R HOOKE T A JEEVES
Direct search solution of numerical and statistical problems
Journal of the Association for Computing Machinery
vol 8 April 1962 pp 212-229

*H2 G HADLEY
Nonlinear and dynamic programming
Addison-Wesley 1964

K1 WILFRED KAPLAN
Advanced calculus
Addison-Wesley 1953

*K2 G A KORN T M KORN
Electronic analog and hybrid computers
McGraw-Hill 1964

K3 W R KLINGMAN D M HIMMELBLAU
Nonlinear programming with the aid of a multiple-gradient summation technique
Journal of the Association for Computing Machinery vol 11 October 1964 pp 400-415

L1 GEORGE LEITMAN editor
Optimization techniques with applications to aerospace systems
Academic Press 1962

M1 K S MILLER F J MURRAY
The mathematical basis for the error analysis of differential analyzers
MIT Journal of Mathematics and Physics nos 2 3 1953

M2 J MUNSON A I RUBIN
Optimization by random search on the analog computer
IRE Transactions EC-8 1959 pp 200-203

*M3 H F MEISSINGER
The use of parameter influence coefficients in computer analysis of dynamic systems
Proceedings of Western Joint Computer Conference May 1960 pp 182-191
SIMULATION vol 3 August 1964 pp 53-63

M4 R A MUGELE
A nonlinear digital optimization program for process control systems
Proceedings Spring Joint Computer Conference vol 21 1962 pp 15-32

*M5 B A MITCHELL JR
A hybrid analog-digital parameter optimizer for ASTRAC II
Proceedings Spring Joint Computer Conference 1964
SIMULATION vol 4 June 1965 pp 399-411

N1 K B NORKIN
On one method of automatic search for the extremum of a function of many variables
Automation and Remote Control (ISA translation) vol 22 May 1961 pp 534-538

*N2 J A NELDER R MEAD
A simplex method for function minimization
Computer Journal vol 7 January 1965 pp 308-313

P1 M J D POWELL
An iterative method for finding stationary values of a function of several variables
Computer Journal vol 5 July 1962 pp 147-151

*P2 M J D POWELL
An efficient method for finding the minimum of a function of several variables without calculating derivatives
Computer Journal vol 7 July 1964 pp 155-162

*P3 M J D POWELL
A method for minimizing a sum of squares of nonlinear functions without calculating derivatives
Computer Journal vol 7 January 1965 pp 303-307

R1 J B ROSEN
The gradient projection method for nonlinear programming
Journal of Society for Industrial and Applied Mathematics
Part I vol 8 1960 pp 181-217 and
Part II vol 9 1961 pp 514-532

*R2 H H ROSENBROCK
An automatic method for finding the greatest or least value of a function
Computer Journal vol 3 October 1960 pp 174-184

R3 L A RASTRIGIN
Extremal control by random scanning
Automation and Remote Control (ISA translation) vol 21 September 1961 pp 891-896

R4 L A RASTRIGIN
The convergence of the random search method in the extremal control of a many-parameter system
Automation and Remote Control (ISA translation) vol 24 November 1963 pp 1337-1342

S1 R I STAKHOVSKII
Twin-channel automatic optimalizer
Automation and Remote Control (ISA translation) vol 19 August 1958 pp 729-740

S2 R I STAKHOVSKII
A multichannel automatic optimizer for solving variational problems
Automation and Remote Control (ISA translation) vol 20 November 1959 pp 1435-1445

S3 A B SHUBIN
10-channel automatic electronic-relay optimizer
Automation and remote control (ISA translation) vol 21 May 1960 pp 432-438

S4 H A SPANG III
A review of minimization techniques for nonlinear functions
SIAM REVIEW vol 4 1962 pp 343-365

*S5 T L SAATY JOSEPH BRAM
Nonlinear mathematics
McGraw-Hill 1964

S6 B V SHAH R J BUEHLER O KEMPTHORNE
Some algorithms for minimizing a function of several variables
Journal of SIAM vol 12 1964 pp 74-92

*T1 R TOMOVIC
Sensitivity analysis of dynamic systems
McGraw-Hill 1963

*W1 HANS S WITSENHAUSEN
Hybrid techniques applied to optimization problems
AFIPS Proceedings Spring Joint Computer Conference 1962 p 377

*W2 DOUGLASS J WILDE
Optimum seeking methods
Prentice-Hall 1964

Z1 G ZOUTENDIJK
Methods of feasible directions
Elsevier 1960